DOCUMENT RESUME

ED 226 722 IR 010 603

AUTHOR Brenner, Lisa P.; And Others

TITLE MTA Computer Based Evaluation System.

INSTITUTION Illinois Univ., Champaign.; Rush Univ., Chicago, IL.

SPONS AGENCY National Institutes of Health (DHEW), Bethesda, Md. Bureau of Health Manpower Education.
PUB DATE 81
GRANT NIH-1-D12-AH-90110
NOTE 18p.; Paper presented at the Annual Meeting of the Association for the Development of Computer-Based Instructional Systems (Atlanta, GA, March 2-5, 1981).
PUB TYPE Reports - Descriptive (141) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Allied Health Occupations Education; Computer Assisted Instruction; *Computer Assisted Testing; *Computer Managed Instruction; Databases; Higher Education; Medical Schools; *Medical Technologists; *Online Systems; Student Evaluation; *Test Items
IDENTIFIERS *Computer Assisted Test Construction; Interactive Systems

ABSTRACT

The MTA PLATO-based evaluation system, which has been implemented by a consortium of schools of medical technology, is designed to be general-purpose, modular, data-driven, and interactive, and to accommodate other national and local item banks. The system provides a comprehensive interactive item-banking system in conjunction with online student management, testing, and analysis or computer grading. Thus a large (up to 16,000) set of test items may be drawn on by specific instructors in creating tests, which may be delivered either online, within the testing component of the system, or printed for off-line delivery. CERL PLATO users and others who have made special arrangements may use the system, providing they have a project and a project director who can access, edit, and direct the project. Existing banks have been developed primarily in the health sciences, but use is not restricted, with applications in medical technology, pharmacology, microbiology, biochemistry, and small projects in other disciplines. The addition of permanent data storage capability will allow an instructor to run an item analysis on a given online test. (LMM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED226722

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☛ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

MTA COMPUTER BASED EVALUATION SYSTEM

Lisa P. Brenner, Ph.D.
Director of Computer Based Education
Rush University

System Design: Lisa Brenner and Gary Czvitkovicz
Programmed by: Gary Czvitkovicz,
Sean Emery, Alan Becker and
Andrew Schweig.

Funded by the National Institute of Health,
Bureau of Health Manpower Grant #1-D12-AH-90110
through the University of Illinois and by
Rush University, Office of Computer Based Education.

Programs and Documentation Copyright, 1979.

Proceedings of the 1981 meeting of the Association
for the Development of Computer Based Instructional
Systems.

IR010603

Introduction

The MTA PLATO based evaluation system provides a comprehensive, interactive, item banking system in conjunction with on-line student management, testing, and analysis. Thus a large (up to 16,000) set of test items may be drawn on by specific instructors in creating tests, which may either be delivered on-line within the testing component of the system, or printed for off-line delivery. The test item, rather than the student or the test itself, is the primary unit of exchange and analysis, and the instructor is provided with maximized resources and flexibility in review, selection, testing, and analysis.

Historically, computers have been used for student evaluation in three areas. The first, item banking, involves the storage and organization of large collections of test items. This application provides instructors with a comprehensive set of evaluation resources (items) to use in creating tests. The computer provides instructors with easy access to these items through search routines which pinpoint questions useful to an instructor. This organization and searching capability provides an important time advantage to an instructor selecting items from a collection of thousands. The MTA evaluation system follows this tradition, and goes on to unite the best of item banking systems with other on-line testing and analysis features.

The second application, CMI, manages student instruction through on-line testing and instructional prescription. CMI systems assume the test as the basic unit. Although management may include a number of applications, the test items themselves are fixed within a single application. The MTA system replicates much of CMI student management; and through its debriefing system can prescribe remediation to students. CMI systems tend to neglect the testing procedure itself, limiting items to simple multiple choice formats. The MTA item bank provides for all question types, from complex matching to patient management problems. The on-line delivery system supports this variety by providing the full range of PLATO display and judging capabilities as a part of its testing component, thus maximizing instructor flexibility and effectiveness in evaluating student performance.

The third area, computer grading, provides test grading and item analysis of given tests, usually in a batch processing mode. This information, which is very useful to evaluating specific test items, is not usually automatically updated in item banks. The MTA system automates this process, while providing thorough on-line or printed test statistics to the instructor.

System Considerations

The MTA evaluation system began as a PLATO based, on-line item banking and testing system for use by a consortium of schools of Medical Technology. Because its primary focus was on the item, the resulting system was based on evaluation flexibility. Practitioners unfamiliar with computers would be using the system, so the second concern was ease of use. The system is heavily prompted and indexed, with only the text and graphics editor requiring any recalled procedures. Members of the consortium had not previously used computers for testing. Thus they submitted the full range of evaluation items, from patient management problems to essays to questions involving complex graphics. The automated testing system was designed to replicate the presentation and judging capabilities of a live instructor. Starting with a data driven question base, the on-line testing system provides the full range of PLATO display and judging options, including graphics, natural language judging, and complex answer analyses for a wide variety of question types. As far as we know, the system is unique in this respect.

The system was designed to be general purpose in order to accommodate other national and local items banks. A complex access system insures high security along with local control. As with all item banking systems, many subscribers would not have on-line access, so the system provides for both on-line and printed items, tests, and analyses.

The system described in this paper is available to CERL PLATO users, and to other PLATO users through special arrangements. Use requires user-provided namesets, one or more to store an item bank (64 word minimum per question), and one to store student management and test information. It also requires project initialization by MTA staff. The CERL PLATO lesson -mtadoc-covers technical specifications for the system.

Existing banks have developed primarily in the Health Sciences, but use is not restricted to that area. The largest current bank is in Medical Technology, including 4,500 items. Other applications cover Pharmacology, Microbiology, and Biochemistry, with small projects in other disciplines. One project has "banked" descriptions of their non-print media collection from which user catalogs are printed. While use of the system is unrestricted, use of any existing bank is by permission from the bank owner/s.

Functional Specifications

This Urbana PLATO-based system was designed as a general purpose, modular, data-driven, interactive system. The components include:

1. A system management program for defining and controlling project access.
2. An editor for creation and display of a hierarchical structure used to organize an item bank.
3. An editor for entering and revising items, associated graphics, judging information, feedback, references, and selection information.
4. A test construction and revision program.
5. A student management system for assigning tests and reviewing test results.
6. An on-line testing system.
7. A data analysis and permanent data storage system. (In progress)
8. Print programs for hard copy of items, tests, answer keys, debriefing information for students, and summary data for instructors.

Each component is a separate interactive PLATO program tied together by an indexing system.

Project Definition and Access

In order to use the MTA system, one must have a "project." The access program allows the MTA system director to create a new project by defining it and attaching the data storage (namesets) necessary for storing evaluation items and tests. The specified project director defines who can access, edit, and direct the project. Figure 1 shows the access choices open to a project director.

A project may have its own item bank and define an independent hierarchy. However, it is also possible to have a project which uses a different project's item bank. In the case of the Medical Technology Association, each school has a project for testing purposes but shares a common item pool. The project director for the common item pool has given "use only" access to each project sharing the bank. No project that has not been given specific access may enter another project's bank. Each project constructs its own set of tests and delivers the tests either within the computer managed instruction component of

the system or by printing the tests for off-line delivery.

Hierarchical Item Organization

Before actually entering the items in the test bank, a hierarchical, tree structure within which the items will be arranged, must be created and defined within the item banking system. The project director has complete discretion in constructing and modifying the hierarchy through the item editor. The press of a function key places the project director into the "add a node mode," where prompting assists in entry and placement. Figure 2 illustrates one page of the structure as it appears on the PLATO screen while in the MTA editor.

Up to 16 major disciplines may be defined and entered by the project director. Each of these disciplines consists of a separately stored "mini-bank" which may contain up to 1000 categorized items along with the item's associated information. Each discipline may be divided into 63 sub-categories, and these sub-categories further divided 63 times. The hierarchy can be extended to 8 levels in depth. An example from the printed nodal structure of categories for the Medical Technology project is illustrated in Figure 3. A given node and everything below it may be moved from one location to another using a separate processing program. Editors and users may go directly to a specified node or may browse through the hierarchy, reviewing or selecting items of interest.

Eventually the bank will allow key-word cross referencing of terms stored with question objectives and by node name. The system may also be modified to allow storage of performance objectives at the node level. (Objectives are currently stored with the item itself.)

The Item Data Base

The item and its descriptive information is stored as data and interpreted by the delivery program. Figure 4 presents a question as stored in the item bank, and Figure 5 lists all information that may be stored with a question.

The node name, stored with a question, can be used for direct access to the question. Items may be located for review or test selection according to one or a combination of prespecified cognitive levels, nodes, skill levels, question types and contributing schools. Although current skill levels have been subjectively determined, performance statistics (difficulty level) may be utilized in the future to automatically adjust this identifier. Other performance and use statistics will be stored in a separate program.

Feedback for any answer consists of up to 300 characters. A given feedback may be associated with a number of answers in a given question. References may also include 300 characters. During the testing process, hints may be made available when a student presses a HELP key. Two hints may be provided for repeated errors with a question. The last date that a question was edited is also stored with each item and questions are tagged as available for instructor-generated tests only or open for self-assessment. Although the system can define up to 16 different variations, the user must indicate which of the question categories listed in Figure 5 are appropriate for the item.

The Editor

Items are interactively entered at the node where they will reside. Initially the editor prompts the person entering the item, asking sequentially for the associated information, objective, context, prompt, answers, feedback, references, and hints. Rule conformity error checking based on question type is done throughout the entry process. Proof-reading for errors of content and typing must be done by hand. Most information associated with the question may be selected by the editor from indexes. A number of default choices such as question type, arrow location, and item status, may be activated or de-activated by the user in order to speed item entry.

A TUTOR-like text and graphics editor allows for easy creation, storage, and revisions of the question itself. Objects may be created, moved in relation to other objects, sized, rotated, and combined with text. The question context may include a number of pages of text/diagrams. The editor "HELP" section provides complete directions and documentation for available commands.

The question text consists of two parts: the context and the prompt. This distinction is made in order to allow the same context (problem description, graphics, or other background information) to be associated with more than one question. In other words, sub-questions are allowed. The editor presents the original, or "used", context on the text editing page for sub-questions. Questions are initially stored through sequential prompting, but later reviews are accomplished by direct access to the item, associated information, answers, feedback, or hints.

Although the same editor is used to store the context for all types of questions, each question type has aspects of the storage and judging process which are unique:

Multiple choice questions include a set of single letter responses each of which may be marked as correct or incorrect. A multiple choice question may consist of one or more acceptable answers or a number of required answers. Feedback may then be associated with one or more of these responses. Alternate response questions are treated as multiple choice with only two allowable choices. The test driver accepts a large number of positive/negative synonyms as appropriate answers for alternate response questions.

Matching questions and chained questions share some features. Both must be treated as one question. Matching questions have the number of selections stored as associated information. A matching question may require one-to-one matches, or one-to-multiple matches. They are treated as a series of multiple choice questions that may be cycled through. Chained questions have previous and subsequent items stored as associated information and may not be used out of context. At this time chained questions do not allow for differential branching. A series of questions chained together may be equated with a patient management problem and may include all types of questions.

Short answer questions may require one or a number of recalled words. Correct and incorrect answers that are entered may establish synonyms and logical relationships between words defined as a single answer, and can allow misspellings and extra words. Numeric answers may include tolerances. Ten distinct words are allowed in a correct/incorrect answer.

Essay questions require a number of sentences. Responses are stored in a PLATO notefile for later on or off-line review and grading by the instructor. No computer judging is involved.

Test Construction

Given a bank of items created with the previously described routines, an instructor may create tests for on-line delivery or printing. Entering the test editor, the instructor is prompted in naming the test and its sections. The instructor selects individual questions or discipline nodes to be included in each section. When selecting individual questions, the instructor may search for items by selecting search criterion from a table of associated information categories. The editor then presents the instructor with those items matching the specified criterion. When the instructor is presented with a satisfactory item it may be selected through a key-press and assigned a scoring weight from 1-10.

Instructors may also select a node and the number of levels down from that node to include as a part of a test section. The system counts the number of items within the specified area that match the selection criteria and presents the tally to the instructor. The instructor selects the number of questions from the total to include in the test. Once a test has been constructed, the instructor may review a list of questions and nodes for each section and may add/delete/revise the items and specifications, such as scoring weight. If the test is to be delivered on-line, different delivery options may be specified for each section of a test. The possibilities are listed in Figure 6.

In setting up delivery options for the on-line test, different approaches to feedback for answers may also be selected. Choices include telling the student if the answer was correct or not (ok,no), presenting the feedback associated with the answer (if matched), and displaying the correct answer if the answer was incorrect. Hints can be made available to students who press the HELP key.

The instructor also specifies the information that will be presented to the student after completion of a test. These debriefing options may either be presented after an on-line test or printed after scoring off-line tests. Debriefing options that may be selected include the question's objective, the question itself, the student's response, the correct answer, feedback for the student's response, and the question reference. (Of course, the instructor must check items to be sure they have objectives, feedback or other information stored with them. If not, a student cannot be presented with the information.)

Tests, once constructed, may be printed for off-line delivery or may be assigned to students using the other management programs. The instructor may review either individual items in a test or an entire test in "student mode" before assigning it.

Student Management System

The student management system allows the instructor to schedule one or more on-line tests to one or more students and to review student performance on the tests. An instructor may copy a PLATO group roster in order to establish a list of students within the management system or may enter them by hand. Each student or templated group of students in the roster may then be assigned up to ten tests. Each test may be given a date and time when it becomes available to the student and a deadline date and time for completion. The instructor may specify that a student will or will not be able to leave the test and finish it

later. For convenience, the instructor is also provided with the ability to add/delete students from a PLATO group.

Once a student has finished a test, the instructor may review the student's responses in relation to the test items and anticipated responses and may change the student score for any item. The review process may be directed to include only items of a specified type (such as short answer), may be comprehensive, or may include only a summary. An item analysis option (in progress) will provide the instructor with a summary of the class performance on the test and an analysis of the the success of each item in predicting test performance.

Future plans will provide a review test option to assign students to a test comprised of all items missed on a previous test; although this option is currently available in a limited way by using the deferred-retry test delivery option.

On-Line Testing System

Once assignments have been made, the student is assigned to the on-line testing program, -mtadrive-. If the student has been scheduled for one or more tests, the program prepares the test, or first presents the student with an index of scheduled tests to select from. Students are automatically warned about delivery criterion specified by the instructor, such as time limits and restarting. The instructor will soon have an instruction page at the beginning of each section which may be used for further directions.

The second program which the student is routed to, handles the process of testing and scoring. Each test section is constructed as it is encountered in order to save space and to insure that current changes to the item bank will be reflected in tests.

The driver follows most PLATO conventions. It uses the standard arrow to indicate a question and requires NEXT's after each answer and between questions. Keys such as SHIFT-HELP to skip a question are advertised. If the question has multiple pages, the student may use the NEXT and BACK keys to review information before answering the question. The driver always allows a student to pass on a question and have it judged incorrect.

The test driver blindly follows the specifications given by the instructor, even if they appear contradictory. For example, an instructor may specify that the student be given the correct answer for missed questions and also that the

student should retry missed questions three times. In this case the driver will give the student the correct answer and then present another arrow for the retry. An instructor may, in fact, use the testing system in the same way as one might use a PLATO lesson or may use it as a blind test with no feedback or assistance.

Permanent Data Storage

This addition to the student management system (in progress) will allow an instructor to run an item analysis on a given on-line test. The results of the analysis will be permanently stored for future reference and will be accessible by item or by test. This process involves the calculation of mean, standard deviation, standard error of measurement for the total test along with the point bi-serial correlations and difficulty level for each item. In order to keep data on a uniform set of students together, the data storage area for this information will be identified within each project by the director. Thus, data on a given population may be kept separate from data on another, and may be reviewed separately.

A cumulative tally of all uses of the item on a test, a cumulative tally of item test performances in the upper and lower third difficulty levels, and an averaged difficulty level of each item, will be stored by the permanent ID number of the item with a backward reference to the current node name. This information will be updated by each item analysis, irrespective of the population. If the text of an item is significantly revised, the summary data for the item may be deleted from within the item editor.

Printing Items, Tests, and Scores

Instructors like to have printed copies of the question bank so that they can construct tests in the office or at home. The MTA system provides for printing each question and its associated information in a format similar to that in Figure 4. The question print program automatically formats items, and prints each out on one or more pages as needed. Each item starts a new page. Prints may be made by discipline, date of entry, date edited, question type or other associated information.

Most of the instructors using the MTA system deliver tests off-line. After they have created a test and set the delivery and debriefing specifications, they can select the test printing program, which takes its information directly from the test editor. The test print program prints the test and section names, the items, and an answer key for the instructor. An instructor may have the test

items scrambled and a number of test versions printed out.

The answer key/s include both selection and debriefing criterion, listing the question number, node, name and correct answer(s).

The debriefing print program utilizes student data from the management system to print indicated information for each missed question. (Information is not printed for questions answered correctly.) Each debriefing form is headed with the name and group of the student and the name of the test.

Print programs give the option of printing directly or storing the output in a dataset. Direct printing assumes a Qume Sprint 5. It also assumes a standard paper size. The indirect method may be used with any printer. Prints can be made at a rate of 55 characters per second when the PLATO system is not busy, however, this is still a relatively slow process.

Conclusions

After three years of development, the MTA Evaluation System is at last reasonably complete and debugged. We look forward to feedback from users and the chance to continue polishing and expanding the system.

Use of computers for testing may not be the most exciting application, but it is the one which most readily appeals to faculty. They can immediately see the improved options for testing students and increased effectiveness. While many of us may prefer developing sophistocated simulations and innovative instructional lessons, the use of PLATO for evaluation can act as a "bread and butter" justification which can lead to continued support for our more esoteric projects.

Figure 1

ACCESS OPTIONS

Change access for: name/group

Current Access: Director

A. Director
B. Author
C. Instructor
D. Proctor
E. Student
F. Self Assessment
G. None

a. Node Hierarchy
b. Questions
c. Tests
d. Test Schedule
e. Test Scores

f. Take Tests
g. Compare Scores
h. See Objectives
i. Take SA Tests
j. Create SA Tests

k. Access List
l. Discipline List
m. Project Info
n. Change Codewords
o. Use Questions

Press a key to change access options

Shift-HELP to delete entry
Shift-NEXT / Shift-BACK to move

Figure 2

Serologic Tests for Syphilis

1	General Principles	(25)*
2	False Positives	(32)*
3	Flocculation Test	(12)*
4	Complement Fixation Tests	(33)*
5	Fluorescent Treponema Antibody (FTA)	(14)*
6	Quality Control	(5)

> 4

Type or touch the number of a subject area and then press NEXT for finer divisions or DATA to add/edit a question at that level.
Or press shift-LAB to add a new subject area.
* denotes bottom node in hierarchy

Figure 3

PRINTED NODE OUTLINE
(One Section of Hematology)

# ques.	node name	node title
0	jdgba	Clinical conditions
1	jdgbb	Laboratory findings
0	jdgc	Hemoglobin D
0	jdgd	Hemoglobin E
11	jdge	Others
1	jdh	Thalassemias
2	jdha	General characteristics
1	jdhb	Beta thalassemia major
0	jdhba	Clinical conditions
2	jdhbb	Laboratory findings
1	jdhc	Beta thalassemia minor
1	jdhca	Clinical conditions
1	jdhcb	Laboratory findings
1	jdhd	Alpha thalassemia major or hemoglobin H disease
0	jdhda	Clinical conditions
0	jdhdb	Laboratory features
0	jdhe	Alpha thalassemia minor
0	jdhf	Delta thalassemia
1	jdhg	Hereditary persistence of feta l hemoglobin
0	jdhh	Hemoglobin lepone
1	je	Leukocytes
2	jea	Origin
5	jeb	Granulocyte formation
3	jeba	Myeloblast

Figure 4

Clinical Significance (Node: jabib04)

ID#: 2424
Date entered: Aug 1979
(contains graphics)

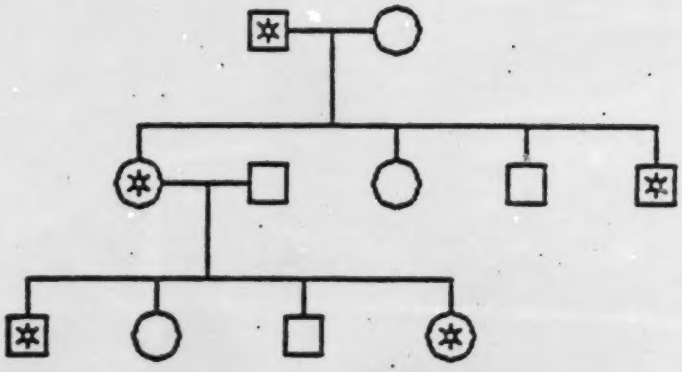
Cognitive level: Application
Skill level: Student
Question type: Multiple choice

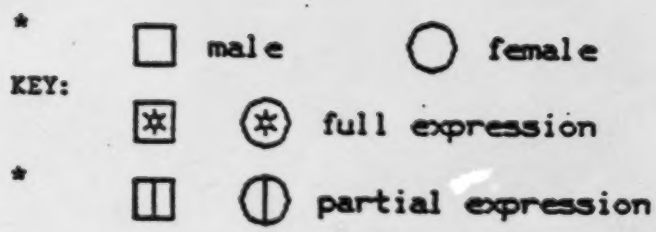
OBJECTIVE:

The student will be able to recognize the genetic disease by studying the genetic pedigree.

QUESTION:

The following is part of a genetic pedigree.





This pedigree is consistent with:

a. alpha thalassemia
b. G-6-PD deficiency
c. hereditary spherocytosis
d. sickle cell anemia

ANSWER: c

FEEDBACK:

REFERENCE: Henry, John B., Clinical Diagnosis and Management by Laboratory Methods, Philadelphia, W.B. Saunders Co., 16th Edition, pp. 986-987, 992, 999-1000, 1004.

Figure 5

Information Included in a Question Record

Question (No limit to length. May include graphics)
Number of Pages
Objective (300 character limit)
Hints (For each consecutive wrong answer)
Feedbacks (For each correct and incorrect answer)
Reference (300 character limit)
Associated Information :
- Node name (defines position in heirarchy)
- Unique ID number (used for permanent data storage)
- Date entered:
- Last changed:
- Last editor:
- Subquestion: (yes or no)
- Context: (If yes, which question?)
- Chain question: (i.e. patient management problem)
  - Previous: (previous question in chain)
  - Next: (next question in chain)
- Question type: (16 possibilities)
  - Multiple choice (all kinds)
  - Matching (all kinds)
  - Short answer (numeric and text)
  - Alternate response (two alternatives)
  - Essay (response stored for instructor review)
- Cognitive level: (8 possibilities)
  - Knowledge
  - Comprehension
  - Application
  - Analysis
  - Synthesis
  - Judgement
- Status:
  - Hold
  - Restricted to instructor tests
  - Open for self-assessment
- Skill level: (8 possible categories)
  - Student
  - Practitioner
  - Specialist
- Contributor: (source of question)

Figure 6

On-Line Delivery Options
For Each Section

I. Presentation and judging options

Time limit -- maximum time student can spend on this section.

Order presented -- show questions in the order listed, or chosen randomly from the list.

Response -- student input can include the touch panel (for questions with touch areas), or be restricted to the keyset. (touch is not operational)

Skip questions -- allow student to skip a question and take it at the end of the section.

Spelling errors -- maximum of 3 misspelled words in short answer questions if allowed.

Extra words -- allow extra words in short answers.

Capitalize -- allow first letter of words in short answers to be capitalized or not capitalized.

II. Response to missed test items

If a question is missed, the student can be given additional chances to answer the question, including:

Continue -- present next question, no additional chances given.

Retry -- allow student to answer question again.

Retest -- present a different question from the same node.

Number of tries -- if retry or retest, the number of additional chances given.

Retake when -- if retry or retest, allow questions to be retaken immediately or at end of section.

III. Student Feedback

Feedback for missed items may include:

Presentation of "ok" or "no" for correct/incorrect

Presentation of feedback associated with a given wrong answer.

Presentation of associated hints if the HELP key is pressed.

IV. Debriefing after the test

After the test is completed, the student may either receive an on-line debriefing or a print which includes the following specified information:

Performance on test compared to class performance (in progress)

Any associated information specified by the instructor, such as objective or feedback.

The question itself.

The student's response as given.